# 1. TITLE

An Automatic Detection of Caries In Dental Radiographs Using Modified Deep Learning Method For Treatment Planning.

# 2. AIMS OF RESEARCH

Dentists frequently do a radiological check on the patient. They can enhance medical treatments by utilizing the deep learning approach in dental images analysis and research to speed up professional judgment. Radiographs help dental professionals and radiologists to monitor, evaluate and diagnose oral diseases and conditions.

Aim of the research mentioned below

1. Automatic caries detection in dental radiograph and condition identification using deep learning approaches which help dentists in their decision-making process and benefit both patients and dental practitioners.
2. Reduce the dentist's workload and assist dentists in their daily practices with precision of dental disease detection and diagnosis.

# 3. INTRODUCTION

A common condition affecting the human oral cavity is dental caries are brought on by bacterial infection. Early identification of the caries' underlying cause will make treatment very simple, which will stop the disease from progressing. Dentists frequently do a radiological check on the patient. They can enhance medical treatments by utilising the deep learning approach in dental images analysis and research to speed up professional judgement. Radiographs help dental professionals and radiologists to monitor, evaluate and diagnose oral diseases and conditions. The Panoramic radiograph is the most commonly used in dental examinations. The different types of dental radiographs are used by the researchers to diagnose oral diseases and conditions. Image segmentation serves as a crucial foundation and is a basic requirement in medical applications for earlier disease diagnosis and condition identification. The image segmentation techniques are boundary, cluster, region, and watershed segmentation. The boundaries of teeth and jaws, may be recognised and marked using semantic segmentation. CNN and Transformer are two semantic segmentation categories based on the backbone network used. UNet powerful architecture used to evaluate biomedical images for several modalities X-rays, CT scans, MRIs, and ultrasounds. U-Net variants are better U-Nets (3D U-Net), attention U-Nets (trims features), inception U-Nets (filters of multiple sizes) and ensemble U-Nets, UNET++, UNET 3+ is an improvement based on U-Net. UNet 3+ has the ability to reduce network parameters in order to enhance computing performance. A network like VNet, FCN, or UNet, 3D UNet, 3D UNet+ResNet for image(2D/3D) segmentation.

## 3.1 Input Modalities:

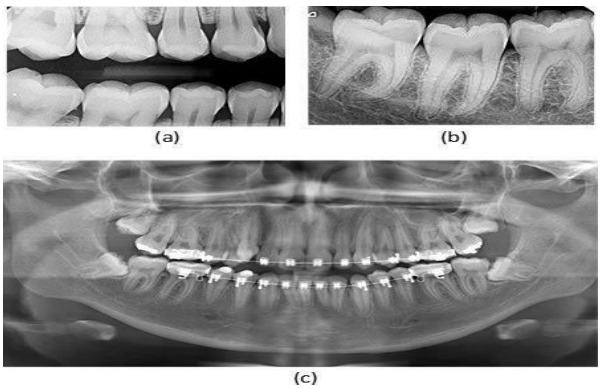


Fig_1: Types of X-ray images: (a) Bitewing X-ray; (b) Periapical X-ray; (c) Panoramic X-ray [28]

Radiographs help dental professionals and radiologists to monitor, evaluate and diagnose oral diseases and conditions. The dental panoramic radiograph is the most commonly used in dental examinations. The different types of dental radiographs are used by the researchers to diagnose oral diseases and conditions; these are mainly of two types: Intra-oral radiography and Extra-oral radiography. The Intra-oral radiography is used to get information of teeth arrangement, root canal infection, and to identify caries. Further this intraoral radiograph is classified as Periapical images (information about the root and surrounding bone regions for 2-3 teeth), Bitewing images (identifying the upper and lower tooth arrangement that are utilised for cap fitting or tooth restoration), Occlusal images (give information about the base of the mouth, displaying the upper or lower jaw's bite). Extra-oral radiography use for full area of the skull and jaws images and classified as Panoramic X-rays ( capture the overall tooth structure used for tooth fractures, trauma, jaw diseases, pathological lesions) shown in Fig_1, Cephalometric X-rays (A view of the head from the side, revealing the teeth, jaw, and surrounding tissues), Sialogram ( x-ray of the salivary ducts and glands), Computed tomography (CT) (3-D interior structures), Cone-beam computed tomography (CBCT) (3D Image for dental structure, nerves, soft tissues and bones). Dental professionals frequently use panoramic radiographs of the jaws because it provides a detailed examination of the patient's maxillary and mandibular regions. The 3D objects are projected into 2D in panoramic images. MSCT is often conducted to assist the diagnosis, treatment, and postoperative efficacy analysis of soft and hard tissue lesions in the maxillofacial region. Although CBCT and MSCT produce 3D images with clear layers, traditional empirical reading could also lead to missed diagnoses and misdiagnoses of early and minor lesions. IOS is generally employed in chair-side digital restoration, digital orthodontics, and digital implant surgery. CBCT and MSCT produce 3D images with clear layers

## 3.2 Deep Learning Approach in Dental Images Analysis and CNN

Recently, convolutional neural networks (CNNs) have achieved great success in image segmentation. In medical image analysis, deep learning approaches are becoming a rapidly growing research area and used to determine if the disease is present or not. Deep learning algorithms' highlighted feature is self-learning. They can train on their own and can be used to both structured and unstructured data sources. Deep learning algorithms listed as CNN-Convolutional Neural Network, LSTM-Long Short-Term Memory Network, RNNs-Recurrent Neural Network, GAN-Generative Adversarial Network, RBF-Radial Basis

Function Network, DBN-Deep Belief Network, MLP-Multilayer Perceptron, SOM-Self Organising Map, RBM-Restricted Boltzmann Machines, Autoencoders. Deep learning advancements have demonstrated the effectiveness of CNNs in handling problems such as image segmentation, classification and object detection. CNN is a powerful algorithm for image processing and the automated processing of images. So it is frequently used for image recognition and classification.

### Convolutional Neural Network (CNN)

A CNN is to identify visual patterns straight from pixel pictures with little to no pre-processing. Image classification is to classify the given image into one of the predefined categories. Feature extraction and classification are common methods for image classification. CNNs extract features using trained convolutional filters, and the network is a fully connected layer that predicts the final label. CNNs have several benefits, including excellent accuracy in image processing and the automated detection of needed features and drawbacks as CNN requires massive labelled data, slower operation.

### Transformer-Based Algorithms

Applying the transformer structure to instance segmentation tasks is a relatively new research area. Cell-DETR was one of the first methods to apply the transformer structure to instance segmentation tasks on biomedical data, achieving performance comparable with that of the latest CNN-based instance segmentation methods, but having a smaller number of parameters and a simpler structure. ISTR, an end-to-end instance segmentation framework, predicted low-dimensional mask embeddings and assigned them with ground-truth mask embeddings to compute the set loss, achieving a significant performance gain for instance segmentation tasks by conducting a recurrent refinement strategy.

## 3.3 Deep Learning Models Task

(Object Detection, Image Classification, Image Segmentation models)

### Object Detection Models:

Medical professionals use object detection and their related points of interest in images as a critical component of diagnosis. The aim is to locate and identify certain tiny parts within a larger picture. These technologies are created to speed up the reading process for human

specialists by automatically detecting areas. The classification and region proposal based detectors are two types of general object detectors. The Region proposal detector with evolution path as RCNN and SPPNet released in 2014→ Fast RCNN and Faster RCNN released in 2015→Mask R-CNN released in 2017→Pyramid Networks/FPN released in 2017→G-RCNN released in 2021. The R-CNN model consists of three modules: region proposal, deep CNN feature extraction, and classification and localization. R-CNN drawback is a high number of redundant regions, multi-step training and time consuming. Many strategies have been proposed to address these issues such as multiple hierarchical segmentation, Edge box method. Fast R-CNN is an upgraded version of R-CNN with faster computation speed and with improved bounding box regression and classification tasks. Mask R-CNN (MRCNN) is with enhanced image and instance segmentation capabilities. MRCNN adds a branch to identify the segmentation mask in each region of interest. A single CNN can predict the bounding boxes and class probabilities. The framework for classification and regression mostly consists of many techniques such as YOLO and SSD released in 2016→RetinaNet released in 2017→YOLOv3 released in 2018→YOLOv4 released in 2020→YOLOR released in 2021. YOLO-You Only Look Once used in real time object detection. YOLO is implemented using the Keras or OpenCV deep learning libraries. YOLO versions are YOLO, YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, and now YOLOv7.

Table_1: Overview on commonly used object detection method for Dental Images

| **Author** | **Method** | **Network** | **Images** |
|---|---|---|---|
| Lawrence Y [1] | CNN | - | Periapical |
| M. Muthu [2] | Deep CNN and GLCM | AlexNet | - |
| Majed B.[3] | ANN and DL | ANN and DL | Periapical |
| Lizheng Liu [4] | MASK R-CNN | ResNet-50 | Clinical Images |
| Devesh Saini [5] | Deep CNN | Inception-V3, Vgg Net16, VggNet19 and Resnet50 | Digital colour image |
| Andac Imak [6] | Deep CNN | AlexNet | Periapical |
| K. Moutselos [7] | DNN Mask R-CNN | - | 88 in-vivo dental images |
| Yusuf Bayraktar [14] | CNN YOLO v3 | DarkNet-53 | Bitewing |

| | | | |
|---|---|---|---|
| Joni Hyttinen [15] | PCA | PCA | Spectral Images |
| Umer Rashid [16] | hybrid mask RCNN | ResNet50 | Periapical |
| Fatemeh S. [17] | CNN-PCA-KNN | | Panoramic |
| Wirtz [18] | CNN | UNet | Panoramic |
| T. Koch [19] | CNN | UNet | Panoramic |
| Jun Zhang [20] | Multi-task FCN | - | CBCT, MSCT |
| Kong Z [21] | CNN | UNet | Panoramic |

According to the research, it was found that InceptionV3, ResNet50, and Xception (extreme inception) took merely 3, 4, and 3 epochs, respectively, to achieve the minimum validation loss. They are able to attain validation accuracy of 99% with a minimal number of epochs. It has been observed that DenseNet-121 and ResNet-50 have the best training accuracy, DenseNet-121 has lowest loss and Inception-ResNet-V2 and VGG16 have greater accuracy for the validation set but not much use for dental images.

### Image Classification:

Image classification involves feature extraction to observe patterns in the dataset. A classical CNN backbone network used for image classification including LeNet, AlexNet, ZFNet, VGGNet, GoogLeNet, Inception, ResNet, DenseNet etc. Network performance judged by Top-1 and Top-5 error rate. Top-1 error rate indicates how many times the network has predicted the correct label with the highest probability. The top-5 number indicates how many times the correct label appears in the network's top five predicted classes.

Table 2: Backbone Network and its variants

| **Network** | **Year** | **Developed By** | **Depth** | **Complexity (GFLOPs)** | **Top-1 Error Rate (%)** | **Top-5 Error Rate (%)** | **No of Parameter (Million)** | **Variants** |
|---|---|---|---|---|---|---|---|---|
| LeNet | 1998 | Yann LeCun et al | 5 | - | - | MNIST: 0.95 | 0.060 | LeNet-5 |
| AlexNet | 2012 | Alex Krizhevsky | 8 | 0.72 | 42.8 | 15.3 | 60 | - |
| ZFNet | 2013 | Matthew Zeiler and Rob Fergus | 8 | | | 11.7 | 60 | - |

| Network | Year | Developed By | Depth | Complexity (GFLOPs) | Top-1 Error Rate (%) | Top-5 Error Rate (%) | No of Parameter (Million) | Variants |
|---|---|---|---|---|---|---|---|---|
| VGGNet | 2014 | Simonyan, Zisserman | 16 | 15.3 | 28.5 | 7.0 | 138 | VGG16/19 |
| GoogleNet | 2014 | Google | 22 | | | 6.67 | 4 | - |
| Inception | 2014 | Google(Szegedy) | 159 | - | 21.2 | ImageNet: 3.5 | 23.6 | Inception (V2, V3, V4) |
| ResNet | 2015 | Kaiming He | 152 | 7.6 | 19.87 | ImageNet: 3.6 | 25.6 | ResNet-18/50/101/110/152/164/200 |
| DenseNet | 2016 | Gao Haung | 190 | - | - | CIFAR-10+: 3.46 | 25.6 | DenseNet-121,161,169,201 and 256 |
| ResNeXt | 2017 | Xie | 10 | 15.6 | 20.4 | CIFAR-10: 3.58 | 68.1 | ResNeXt-50/101/152 |
| Xception | 2017 | Francois Chalet | 71 | 11 | 21 | 5.5 | 22.9 | Xception, Xception-41, 65,71 |
| EfficientNet | 2020 | Tan | 237-813 | 0.39, 37 | 22.9, 15.7 | 0.788 | 11 | EfficientNet B0 to B7 |

The CNN-based architectures have been evaluated using generally three metrics including Top-1 error, Top-5 error, and the computational complexity of the model using the number of Flops. These metrics are used for comparing the effectiveness of the proposed models presented in Table 2. EfficientNet-B7 have reached minimum values of Top-1 and Top-5 error rates compared with Inception-v3 that has the second-best results.

### Semantic Segmentation:

CNN and Transformer are two semantic segmentation categories based on the backbone network used. A network like VNet, FCN, or UNet, 3D UNet, 3D UNet+ResNet for image(2D/3D) segmentation. The fully convolutional network model (FCN) is a typical representative network that adopts convolutional layers in place of the original fully connected layers to make pixel-level predictions for semantic segmentation. Later, an end-to-end segmentation model named U-Net was proposed for medical image segmentation by Ronneberger et al. UNet is a powerful method used to evaluate biomedical images for several modalities X-rays, CT scans, MRIs, and ultrasounds. It has an encoder-decoder

structure, similar to a U shape. In view of the superior performance of U-Net, several variants were presented including U-Net variants are better U-Nets(3D U-Net), attention U-Nets (trims features), inception U-Nets (filters of multiple sizes) and ensemble U-Nets, UNET++, UNET 3+ is an improvement based on U-Net. UNet 3+ has the ability to reduce network parameters in order to enhance computing performance.

## Backbone Framework:

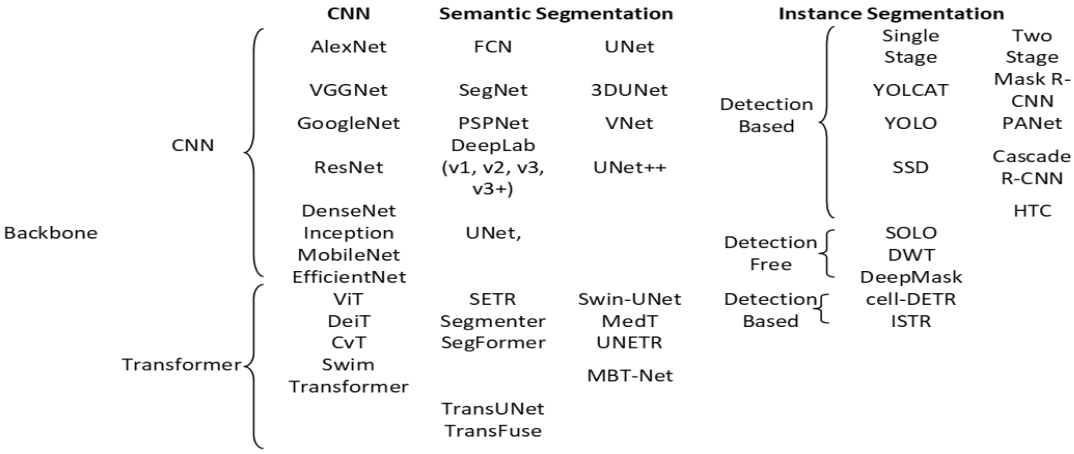


Fig_2: Backbone Framework [27]

For the backbone network, there are CNN-based and Transformer-based methods, the former includes AlexNet, VGG, GoogLeNet, ResNet, DenseNet, MobileNet, ShuffleNet, and EfficientNet, and the latter includes ViT, Data-efficient image Transformers (DeiT), Convolutional vision Transformer (CvT), and Swim-Transformer.

For the semantic segmentation, the CNN-based methods include FCN, SegNet, PSPNet, DeepLab (v1, v2, v3, v3+), UNet, VNet, and UNet++, and the Transformer-based methods include SETR, Segmenter, SegFormer, Swin-UNet, Medical Transformer (MedT), UNETR, MBT-Net, TransUNet, and TransFuse.

The instance segmentation task also can be categorized into CNN-based and Transformer-based methods.It can be divided into the detection-based and the detection-free instance segmentation methods, the former is divided into the single-stage (YOLCAT, YOLO, and SSD) and two-stage methods (Mask R-CNN, PANet, Cascade R-CNN, and HTC), and the latter includes SOLO, DWT, and DeepMask. The Transformer-based methods, such as cell-DETR, ISTR, belong to detection-based methods.

# 4. LITERATURE SURVEY

The literature review involves a comprehensive review of existing research and studies related to dental disease detection, deep learning, and segmentation algorithms applied to various radiographs. The literature review serves as the foundation for the current research by highlighting the gaps in knowledge, identifying previous approaches, and showcasing the need for further investigation. In the research, pre-trained deep models have been used commonly, besides traditional feature extraction methods. Further literature studies based on the tooth detection, dental implant, filled teeth, caries detection commonly for panoramic and periapical imaging systems have been used, hybrid models and segmentation method.

### Tooth Detection: Dental Decay, Dental Implant, Filled Teeth, Caries Detection

Lawrence Y et al [1], they proposed a CNN method, a 18 layer architecture with 7 experiments to classify whether a normal tooth, implants, fillings, teeth cavity. The dataset included a total of 3953 dental X-rays. Used CNN with keras python with 18-layer architecture. A total of seven experiments were done, each with four distinct epoch counts. The outcome of their project showed 93.04% average accuracy, accuracy rate is 87 % for normal teeth, 98% for implants and fillings and 89% for cavities.

M. Muthu Lakshmi et al [2], A total of 1700 images were taken, with 900 utilized for training and 800 for testing. Deep CNN is employed for dental image prediction. The deep CNN algorithm in the AlexNet package achieves 98.60 % accuracy.

Majed Bouchahma et al [3], Using ANN and DL approaches, an automatic method for detecting deterioration from dental X-ray pictures is developed. This method predicts, filling, root canal treatments and fluoride. 200 periapical X-ray images were used.They utilized the Keras API in Python to build the system. Method showed 87% overall accuracy, The total success rate was 98% for fluoride detection, 88% root canal detection, 77% filling treatment.

Lizheng Liu et al [4], A smart dental device intended to identify dental disorders utilizes 600 clinical images to train the model for auto diagnosis. The MASK R-CNN model generates results with up to 90% accuracy, as well as excellent sensitivity and specificity.

Devesh Saini [5], Proposed Deep CNN for dental caries detection. With a dataset of 500 color teeth images, of which 400 were utilized for training, 70 for validation, and 30 for testing, they compared the performance of the VGGNet16, VGGNet19, Resnet50, and Inception-V3 networks. Results indicate that the inception-V3 beat other models such as VGGNet16, VGGNet19, and Resnet50 for the dataset with 99.89% accuracy and loss of 0.01%.

Yusuf Bayraktar et at [14], They used YOLO v3 to detect dental caries showed result with 13.89% lesion frequency for test data and 94.59% overall accuracy and parameter like perceptivity, PPV, NPV and Overall AUC was assessed in bitewing radiographs for premolars and molars with good results for detecting dental caries.

Joni Hyttinen [15], Recession of the gingiva, gingivitis, calculus, caries, granules of Fordyce, leukoplakia, and pigmented lesions was demonstrated in spectral image-based visibility enhancement. 116 Spectral images used in analysis. PCA (Principal Component Analysis) derived computational filters compared with Michelson contrasts are improved in all the investigated cases.

### Hybrid Models

Umer Rashid et at [16], The ResNet50 was used as a pre-trained network. A hybrid mask RCNN model showed 78% accuracy and 92 %, labelled datasets correctness can reach 96% using periapical radiographs 508 imaged for training and 100 for testing and validation. The evaluation metrics were average recall and precision, OKS (Object Key point Similarity).

Fatemeh S. et al. [17], Current age estimation approaches rely significantly on time-consuming and manual image processing. They suggested HCNN-KNN as Hybrid models combine CNN and KNN for estimating the age of teeth. Proposed CNN-PCA-KNN model as CNN for Feature extraction, PCA for Over fitting and KNN for classifier. The 15 to 23 years old child 1,922 panoramic dental radiographs are used with 1,3,6 month and a year interval resulting in accuracies of 99.98, 99.96, 99.87, and 98.78. Here are datasets from diverse places, with distinct features, and varying age ranges. When tested and cross-validated, the proposed model outperforms and faster than existing models such as CNN, ResNet, and GoogLeNet Inception.

## UNet for dental image segmentation

U-net is an image segmentation technique developed primarily for image segmentation tasks. These traits provide U-net with a high utility within the medical imaging community and have resulted in extensive adoption of U-net as the primary tool for segmentation tasks in medical imaging. The success of U-net is evident in its widespread use in nearly all major image modalities, from CT scans and MRI to X-rays and microscopy. Furthermore, while U-net is largely a segmentation tool, there have been instances of the use of U-net in other applications. Given that U-net's potential is still increasing, this narrative literature review examines the numerous developments and breakthroughs in the U-net architecture and provides observations on recent trends

Kong et al [21] created a large maxillofacial dataset and proposed an efficient encoder-decoder network model named EED-Net to solve this problem. In order to obtain more semantic features from the depth and breadth, they replace the convolution layer with the residual block in the encoder and adopt Inception-ResNet block in the multipath feature extractor. A weighted loss function is used to enhance segmentation accuracy. The comprehensive experimental results on the new dataset demonstrate that their model achieved accuracy is 99.28% and speed trade-offs for maxillofacial segmentation than the latest methods.

Lee et at [22] used a histogram-based method as a preprocessing step to estimate the average gray density level of the bone and tooth regions. Also, we developed a posterior probability function. They regularize the CNN models with spatial dropout layers and replacing the convolutional layers with dense convolution blocks further improved the segmentation performance.

N. Siddique [25], the researcher reviewed the different image modalities and application areas that have been enhanced by U-net. They explored the many variants of U-net and its diverse applications on a multitude of image modalities. Also examined the major deep learning methods and their application areas for all of the papers in their survey.

Table_3: Performance-wise survey table for deep learning model for dental image analysis

| **Author** | **Method** | **Network** | **Images Type** | **No of Images** | **Database Publicly availability** | **No of Class _Class (Disease)** | **Metrics** | **Accuracy** |
|---|---|---|---|---|---|---|---|---|
| Lawrence Y [1] | CNN | - | Periapical (top to bottom jaw) | 3953 | No | 4_Normal teeth, implants, fillings, teeth cavity. | Average accuracy | Avg accuracy 93.04% |
| M. Muthu [2] | Deep CNN and GLCM | AlexNet | Dental Images | 1700 | No | 1_Dental illness | Precision, Recall, Accuracy and F1 Score | Overall Accuracy 98.60% |
| Majed B [3] | ANN and DL | ANN and DL | Periapical | 200 | No | 3_Fluoride, Filling, Root canal | Success rate | Overall Accuracy 87%. |
| Lizheng Liu [4] | MASK R-CNN | ResNet-50 | Clinical Images | 12,600 | No | 7_Periodontal disease, Dental fluorosis, Decayed tooth, Broken tooth, Calculus, Plaque, Caries | Accuracy, Sensitivity, Specificity | Accuracy dental caries (90.1%) |
| Devesh Saini [5] | Deep CNN | Inception-V3, Vgg Net16, VggNet19 and Resnet50 | Digital colour image | 500 | No | 1_Dental Caries | Accuracy, Loss | For Inception-V3 99.89% accuracy and loss of 0.01% |

| Author | Method | Network | Images Type | No of Images | Database Publicly availability | No of Class _Class (Disease) | Metrics | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Andac Imak [6] | Multi Input Deep CNN | AlexNet | Periapical | 340 | No | 1_Dental Caries | Accuracy | 99.13%. |
| K. Moutselos [7] | DNN Mask R-CNN | - | In-vivo dental images | 88 | No | 1_Occlusal Caries | micro-F-measure for the mc, cpc, wc | 0.889, 0.778, 0.667. |
| Mukul Goswami [8] | CNN | - | Colour images | 598 | No | 2_Oral Cancer and Dental Caries | Accuracy | 83-94% |
| Yusuf Bayraktar [14] | CNN YOLO v3 | DarkNet-53 | Bitewing | 1000 | No | 2_Caries, lesions | Accuracy, Sensitivity, Specificity | overall accuracy 94.59% |
| Joni Hyttinen [15] | PCA | PCA | Spectral Images | 316 | yes | 1_Oral lesions | c: Michelson contrasts, F: filter, cp: Penalized contrast scores | - |
| Umer Rashid [16] | hybrid mask RCNN | ResNet50 | Periapical | 508 | No | 1_Dental caries | Avg precision, avg recall, and OKS | accuracy is 78% and 92 % |

| Author | Method | Network | Images Type | No of Images | Database Publicly availability | No of Class _Class (Disease) | Metrics | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Fatemeh S. [17] | CNN-PCA -KNN | - | Panoramic | 1922 | No | 1_Age estimation using OPG | Accuracy | 99.98% |
| Wirtz [18] | CNN | UNet | Panoramic | 24 | No | 1_Jaw Segmentation | Accuracy, Precision, Recall, Specificity, F1-Score | 0.818 |
| T. Koch [19] | CNN | UNet | Panoramic | 1500 | No | 1_Jaw Segmentation | DSC | 0.934 |
| Jun Zhang [20] | Multi-task FCN | - | CBCT, MSCT | 107 | No | 1_Joint craniomaxillofacial bone segmentation | DSC, SEN, PPV | - |
| Kong Z [21] | CNN | UNet (EED-Net) | Panoramic | 2602 | No | 1_Maxillofacial segmentation | Accuracy, Jaccard, HD, PPS | 0.9928 |
| Lee[22] | CNN | UDS-Net (U-Net + dense block + spatial dropout) | CBCT | 102 | No | 1_Dental Implant | DSC_Val, DSC_Test | 0.935, 0.915 |

## Discussion:

Deep learning algorithms help in the rapid diagnosis of diseases and dental issues. CNN has a strong learning ability, and can automatically learn the important features of images. As per our observation, accurate results are observed in pre-trained DL models as F-RCNN, M-RCNN, VGG, YOLO. RCNN is good for teeth segmentation for diagnosis and forensic identification. M-RCNN is used to identify tooth regions (predict multiple ROI) via multiple image types compared to F-RCNN. The YOLO algorithm has a limitation in that it struggles with small objects inside the image, although it performed well in bitewing dental images [14]. Also observed that frequently used deep learning models are VGG16, VGG19, ResNet50, ResNet101, DenseNet121, and InceptionV3; their fusion may produce magical results. For dental image segmentation ResNet101, EfficientNet and DenseNet121 backbone network is used to achieve targeted accuracy.

Image segmentation methods have been improved over the past decades as segmentation being better than detection is that it is learning to identify each pixel as part of the object. Majorly used semantic segmentation is UNet and Mask RCNN for instance segmentation in medical image segmentation, but it remains a complex and difficult process due to the dissimilarity in the images. Research has been done on the segmentation of dental X-rays. However, the analysis of dental images has some difficulty in comparison with other medical images. According to a survey, there were some issues with dental image segmentation as given below:

- Irregular shape of object/structure/topology
- Intensity variations in X-Ray,
- Limitation of capturing devices
- Quality caused by conditions such as low contrast and noise (Irregular and blurred contours)
- Hybrid dataset

Due to all these problems, it is always difficult to find an accurate and appropriate segmentation method for all types of dental X-ray images. Also, deep learning models have the problem that the input and output of the network are well understood, but the behaviour of the internal hidden layers is not. This creates a problem in which researchers often do not understand how to fix errors in the network or which layers or filters are more important to the task. Other implementation challenges associated with Hyperparameter tuning, Execution time, Generalization and Precision compromising. Also, complexity to identify small multiple objects in an image and dental structure varies from person to person. There are few challenges associated specifically for the dental image segmentation as mentioned below

- Complexity in determining the orientation of teeth structure as size and shape varies person to person
- Variation of individual tooth sizes per person
- Location and orientation of teeth
- Space of missing teeth or gaps in two teeth
- Root canal treatment diagnosis
- Lack of availability of large datasets
- Presence of superfluous teeth in somewhat rare cases; leading to unsuccessful segmentation.
- Find the accurate position for the placement of dental implants
- Shape, number, and position of teeth are the main targets of a dentist when screening for a patient's problems on X-rays.

## RESEARCH GAP

- Automatic and accurate tooth detection, segmentation remains a challenging task as individual teeth size varies with different dental illnesses.
- In the survey, very few papers on the multiclass classification and majority of all the papers on the single class or binary class classification such as caries detection or specific dental illness detection based. I would like to propose a method with multiclass/ Multilabel classification.
- Two to three papers are on the number of teeth prediction, but no naming and labeling permanent teeth under universal numbering system in the image, I would like to propose a method for teeth numbering with labels and total tooth count using DL approach
- Processing the hybrid dataset, 2D or 3D images is one of the biggest challenges in medical image analysis. The majority of the papers are on one type of modalities as well as two dimensional radiographs. Very few papers on hybrid datasets or three-dimensional radiographs.

# 4. OBJECTIVES OF RESEARCH

The objectives of a research on dental disease detection using deep learning algorithm performance for various radiographs. Below are some key objectives that a research might pursue:

1. To select and modify a deep learning model suitable for the detection of dental caries in panoramic dental images
2. To develop and optimize a deep learning model for accurate detection and localization of dental anomalies, such as caries (tooth decay), in dental images
3. To implement a comprehensive training, validation, and testing framework for analyzing dental cavities, enabling automated decision-making in dental diagnostics.
4. Performance evaluation and comparison with Dentists/Expert judgement (before vs. after revision) with assistance of the DL model.

## SIGNIFICANCE OF THE RESEARCH:

The significance of the research lies in its potential contributions to the field of dental radiography and disease detection. Some of the key aspects of the research significance include:

- This can lead to earlier detection and intervention, potentially improving patient oral health and reducing treatment costs.
- Save valuable time for dental practitioners by assisting in the analysis of radiographs.
- Automate decision-making processes and enhance their workflow by integrating the developed deep learning algorithm into dental clinics
- It can be applicable in diverse clinical scenarios and improve the analysis of different dental conditions.

## INTERNATIONAL AND NATIONAL STATUS

In this technology-intensive world, individuals, businesses, and governments use Artificial Intelligence (AI) to automate their workflows and minimize redundant tasks. The research would consider the real-world clinical applicability of the deep learning algorithm. Factors such as processing time, ease of integration into existing dental software, and compatibility with different radiography systems would be evaluated. The clinical trials explore the feasibility and effectiveness of DL method in detecting various dental diseases, including caries, periodontal disease, and other dental abnormalities with regulatory considerations and adherence to patient data privacy, data protection, and informed consent when using medical data for algorithm training and validation.

International organizations, such as the World Health Organization (WHO) and dental associations, and national organizations such as National Dental Association (NDA) may be interested in automatically detecting disease or abnormalities from X-ray images and classifying them into several disease types or severities in radiology. As DL technologies continue to evolve, the international status may pave the way for establishing guidelines, standards, and ethical considerations regarding the integration of deep learning algorithms in dental healthcare.

## RESEARCH HYPOTHESES:

- Development of DL-based dental disease detection system for early diagnosis, monitoring, treatment planning, and preventive care to improve oral health
- To review the conventional methods for detecting dental disorders using digital dental radiographs

# 5. METHODOLOGY FOR CLASSIFICATION, DETECTION, LOCALIZATION DENTAL DISEASE

The proposed deep learning method can be used for classification, detection, localization on large dental images dataset. The proposed method used to classify dental diseases in dental images to provide diagnostic data about dental issues, abnormalities, and conditions, trained using the input radiographs. Dental disease areas and structures of teeth are the respective target regions. The methodology for Dental disease detection using deep learning segmentation algorithm performance for various radiographs involves a systematic and comprehensive approach to evaluating and comparing the efficacy of different deep learning algorithms in dental disease detection across multiple radiographic modalities. The proposed method performance can be compared with the traditional methods. The Steps are as follow

1. ***Data Collection:*** The collection of a diverse and representative dental radiographs dataset encompasses a wide range of dental diseases, including caries, periodontal conditions, and other abnormalities, along with normal cases for comparison.

2. ***Preprocessing:*** The collected radiographic images undergo preprocessing steps, such as resizing, normalization, and noise reduction, to ensure consistency and improve the quality of the data. Data augmentation techniques are applied to augment the dataset.

3. ***Model Selection:*** The deep learning segmentation algorithms are considered for dental disease detection. These models are chosen based on their track record in medical image segmentation tasks and suitability for handling diverse dental radiographs.

   The modified UNet segmentation model is mostly used to accurately segment individual teeth from dental images. This involves training the model to identify the boundaries of each tooth and generate precise segmentation masks.

### Attention U-net

An often-desirable trait in an image processing network is the ability to focus on specific objects that are of importance while ignoring unnecessary areas. The attention U-net achieves this by making use of the attention gate [30], [31]. An attention gate is a unit which trims features that are not relevant to the ongoing task. Each layer in the expansive

path has an attention gate through which the corresponding features from the contracting path must pass through before the features are concatenated with the upsampled features in the expansive path. Repeated uses of the attention gate after each layer improves segmentation performance significantly without adding too much computational complexity to the model.

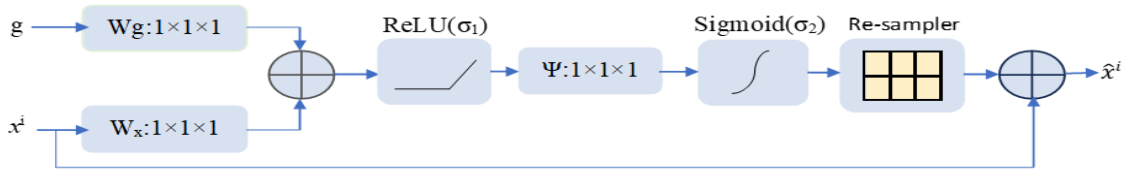


Figure 3. Additive attention gate schematic.

The input signal $x^i$ and the gating signal g both pass through separate 1x1x1 convolutions. The signals are then added and undergo a series of linear transformations which are ReLU activation (σ1), a 1x1x1 convolution, sigmoid activation (σ2), and an optional grid resampler. Finally, the original input is concatenated to the output from the sigmoid unit or the resampler.

The attention unit is useful in encoder-decoder models such as the U-net since it can provide localized classification information as opposed to global classification. In U-net this allows different parts of the network to focus on segmenting different objects. Furthermore, with properly labeled training data, the network can attune to particular objects in an image. The attention gate applies a function where the feature map is weighted according to each class and the network can be tuned to focus on a particular class [33], and hence pay attention to particular objects in an image. While there are different types of attention gates, additive attention is more popular in image processing due to it resulting in higher accuracy. Attention U-net has been applied on problems such as ocular disease diagnosis [32]–[37], melanoma [38], lung cancer [33], cervical cancer [39], abdominal structure segmentation [31], fetus development [31], and brain tissue quantification[40].

I would like to use modified attention UNet for dental caries detection and classification.

4. ***Training and Validation:*** The selected deep learning models are trained on the preprocessed dataset using appropriate loss functions and optimization algorithms. The training process involves splitting the dataset into training and validation sets to monitor the model's performance and prevent overfitting.

   Ex: The trained modified UNet model can identify and classify various dental conditions such as dental caries, periodontal disease allowing for targeted treatment and

intervention.To enhance the diagnostic capabilities, the UNet segmentation model can be extended to perform multi-class segmentation.

5. ***Performance Evaluation:*** The trained deep learning models are evaluated using various performance metrics, such as accuracy, sensitivity, specificity, Dice coefficient, and Intersection over Union (IoU). The evaluation is conducted on separate test datasets to assess the models' ability to accurately segment dental diseases in different radiographic modalities.
6. ***Comparison and Analysis:*** The results of each deep learning model are analyzed and compared to identify the most effective algorithm for dental disease detection. The strengths and weaknesses of each model are assessed, considering factors like computation time, complexity, and ease of implementation.

## Dataset

Dataset is obtained from Dr Narkhede Dental Clinic, Thane. The data set consists of big number of dental panoramic X-ray images, including 70% images for training, 30% images for validation and testing

**Framework for dental disease detection and classification**

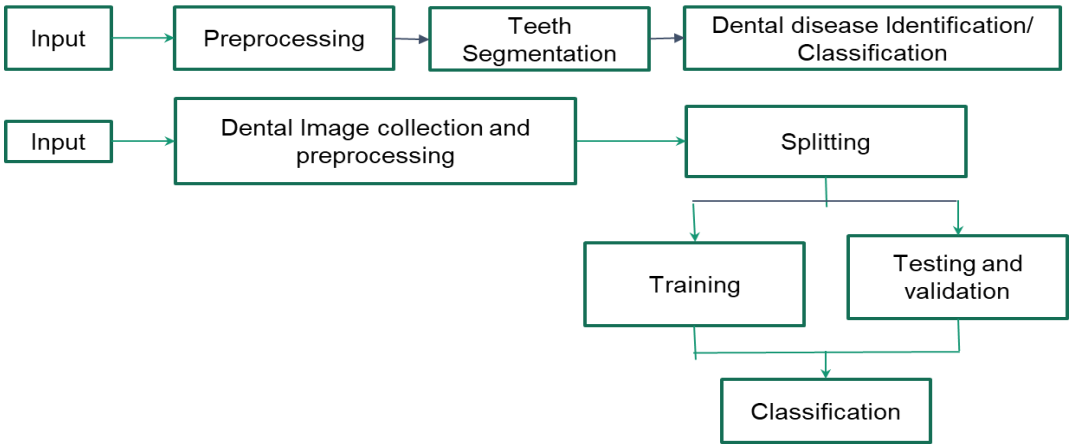


Fig_4: a) Framework for dental disease detection and classification, a) dataset processing flow

## Expected outcomes of research:

Detection, localization, classification and analysis of dental diseases and conditions and other dental abnormalities using deep learning approaches in dental radiograph analysis which help dentists in their decision-making process and benefit dental practitioners for accurate and precise diagnosis and report generation

# 6. STATISTICAL PLAN

Table_4: Statistical Plan

| Year | Year1 | Year 2 | | Year 3 | |
|---|---|---|---|---|---|
| | | **Odd Semester** | **Even Semester** | **Odd Semester** | **Even Semester** |
| **Objectives** | Course Work | Data collection and Data preprocessing | Suitable Model Selection | To accurately and precisely detect and localize of caries (tooth decay) within dental images, | Performance evaluation and comparisons of dentists (before vs. after revision) with assistance of the DL model. |
| **Target** | | Data collection and Data preprocessing | Modify the training model | Training, testing and validation the DL Model | Performance evaluation and comparisons |
| **Expected Outcome** | | Input Dataset for training and Testing | Model development | Trained model, Results and Comparison | Assess the practical applicability |
| | | **Paper Publication** | | **Paper Publication** | |

# 7. LIMITATIONS OF THE RESEARCH:

Every research has its limitations that may affect the interpretation and generalizability of the findings. For the research on dental disease detection using deep learning segmentation algorithm performance for various radiographs, some potential limitations include:

- Dental radiographs can exhibit variations due to different imaging devices, patient positioning, and image quality. Such variations may affect the algorithm's accuracy and reliability.
- The manual annotations provided by dental experts may have inherent inter-observer variability. This variability could introduce uncertainty in the ground truth used for algorithm training.

- **Complexity in determining the orientation of teeth structure**: The size, shape, and location of teeth and dental disease are complex and vary from person to person. Find the accurate position for the placement of dental implants

Also according to the literature research some other limitations are 1) Class imbalance in the dataset can impact the algorithm's performance, especially on detecting less prevalent dental conditions, 2) Deep learning models can be computationally demanding, requiring substantial processing power and memory. Limited computational resources may restrict the model's complexity and performance, 3) Implementing the algorithm in real-time clinical settings may pose challenges due to processing time constraints. The algorithm's speed and efficiency need to be considered for practical application.
In my research, I will try to attempt the above few limitations.

# 8. CONCLUSION

Deep CNN algorithms, according to the findings of this research, provide comprehensive, reliable, and accurate image evaluation and disease identification, enabling rapid and effective diagnosis that are analogous to dental professionals and radiologists' opinions. The research will focus on accurate and precise detection and localizing dental anomalies, such as caries (tooth decay), periodontal disease, and other abnormalities, within dental images.
Image segmentation methods have shown increased model performance for medical images. The research will focus on automatic segmentation and classification of individual teeth by its label, according to diseases through dental images, along with analysis of dental structures.

# 9. REFERENCES


[1] Lawrence Y. Deng, See Sang Ho, Xiang Yann Lim, “Diseases Classification Utilizing Tooth X-ray Images Based On Convolutional Neural Network”, 2020 International Symposium on Computer, Consumer and Control (IS3C), Conference Paper, Publisher: IEEE

[2] M. Muthu Lakshmi, P. Chitra, “Tooth Decay Prediction and Classification from X-Ray Images using Deep CNN”,2020 International Conference on Communication and Signal Processing (ICCSP), Conference Paper, Publisher: IEEE

[3] Majed Bouchahma et al, “An Automatic Dental Decay Treatment Prediction using a Deep Convolutional Neural Network on X-Ray Images’, 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA), Conference Paper, Publisher: IEEE

[4] Lizheng Liu et al, “A Smart Dental Health-IoT Platform Based on Intelligent Hardware, Deep Learning, and Mobile Terminal”, IEEE Journal of Biomedical and Health Informatics, Volume: 24, Issue: 3, Journal Article, Publisher: IEEE

[5] Devesh Saini, Richa Jain, Anita Thakur, “Dental Caries early detection using Convolutional Neural Network for Tele dentistry”, 7th International Conference on Advanced Computing and Communication Systems (ICACCS-2021), Volume: 1, IEEE Conference.

[6] A. Imak, A. Celebi, K. Siddique, M. Turkoglu, A. Sengur and I. Salam, "Dental Caries Detection Using Score-Based Multi-Input Deep Convolutional Neural Network," in IEEE Access, vol. 10, pp. 18320-18329, 2022, doi: 10.1109/ACCESS.2022.3150358.

[7] K. Moutselos et al, “Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning”, 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Year: 2019, IEEE

[8] Mukul Goswami et al, “Automated Detection of Oral Cancer and Dental Caries Using Convolutional Neural Network”, 9th ICRITO, Year: 2021, Publisher: IEEE

[9] Joni Hyttinen et al, “Oral and Dental Spectral Image Database—ODSI-DB” *Appl. Sci.* 2020, *10*(20), 7246.

[10] Karen Panetta et al, “Tufts Dental Database: A Multimodal Panoramic X-Ray Dataset for Benchmarking Diagnostic Systems”, IEEE journal of biomedical and health informatics, vol. 26, no. 4, April 2022.

[11] Rad, A.E., Rahim, M.S.M., Rehman, A. et al. Digital Dental X-ray Database for Caries Screening. 3D Res 7, 18 (2016). https://doi.org/10.1007/s13319-016-0096-5

[12] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon and L. Oliveira, "Deep Instance Segmentation of Teeth in Panoramic X-Ray Images," 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, Brazil, 2018, pp. 400-407, doi: 10.1109/SIBGRAPI.2018.00058.

[13] A. H. Abdi, S. Kasaei, and M. Mehdizadeh, “Automatic segmentation of mandibles in panoramic X-ray,” J. Med. Imag., vol. 2, no. 4, 2015, Art. no. 044003

[14] Yusuf Bayraktar, Enes Aya, “Diagnosis of interproximal caries lesions with deep convolutional neural network in digital bitewing radiographs” PMID: 34173051 PMCID: PMC8232993 DOI: 10.1007/s00784-021-04040-1, Free PMC article

[15] J. Hyttinen, P. Fält, H. Jäsberg, A. Kullaa and M. Hauta-Kasari, "Computational Filters for Dental and Oral Lesion Visualization in Spectral Images," in IEEE Access, vol. 9, pp. 145148-145160, 2021, doi: 10.1109/ACCESS.2021.3121815.

[16] Rashid U, Javid A, Khan AR, Liu L, Ahmed A, Khalid O, Saleem K, Meraj S, Iqbal U, Nawaz R. A hybrid mask RCNN-based tool to localize dental cavities from real-time mixed photographic images. PeerJ Comput Sci. 2022 Feb 18;8:e888. doi: 10.7717/peerj-cs.888. PMID: 35494840; PMCID: PMC9044255.

[17] Fatemeh S. et at, “Hybrid HCNN-KNN Model Enhances Age Estimation Accuracy in Orthopantomography”, Digital Public Health, 30 May 2022.

[18] Wirtz A, Mirashi SG, Wesarg S. “Automatic teeth segmentation in panoramic X-ray images using a coupled shape model in combination with a neural network. In: International conference on medical image computing and computer-assisted intervention” Cham: Springer (2018). p. 712–9.

[19] Koch TL, Perslev M, Igel C, Brandt SS. “Accurate segmentation of dental panoramic radiographs with U-Nets. In: 2019 IEEE 16th International Symposium on Biomedical Imaging”, (ISBI 2019). Venice: IEEE (2019). p. 15–9. doi: 10.1109/ISBI.2019.8759563

[20] Jhang J et al, “Joint craniomaxillofacial bone segmentation and landmark digitization by context-guided fully convolutional networks” In: International conference on medical image computing and computer-assisted intervention. Cham: Springer (2017).p. 720–8

[21] Z. Kong et al., "Automated Maxillofacial Segmentation in Panoramic Dental X-Ray Images Using an Efficient Encoder-Decoder Network," in IEEE Access, vol. 8, pp. 207822-207833, 2020, doi: 10.1109/ACCESS.2020.3037677.

[22] S. Lee, S. Woo, J. Yu, J. Seo, J. Lee and C. Lee, "Automated CNN-Based Tooth Segmentation in Cone-Beam CT for Dental Implant Planning," in IEEE Access, vol. 8, pp. 50507-50518, 2020, doi: 10.1109/ACCESS.2020.2975826.

[23] Fernandez, Kelwin, and Carolina Chang. "Teeth/palate and interdental segmentation using artificial neural networks." IAPR Workshop on Artificial Neural Networks in Pattern Recognition. Springer, Berlin, Heidelberg, 2012.
[24] S. -Y. Lin and H. -Y. Chang, "Tooth Numbering and Condition Recognition on Dental Panoramic Radiograph Images Using CNNs," in *IEEE Access*, vol. 9, pp. 166008-166026, 2021, doi: 10.1109/ACCESS.2021.3136026.
[25] N. Siddique, S. Paheding, C. P. Elkin and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," in IEEE Access, vol. 9, pp. 82031-82057, 2021, doi: 10.1109/ACCESS.2021.3086020.
[26] S. -L. Chen et al., "Missing Teeth and Restoration Detection Using Dental Panoramic Radiography Based on Transfer Learning With CNNs," in IEEE Access, vol. 10, pp. 118654-118664, 2022, doi: 10.1109/ACCESS.2022.3220335.
[27] G. Silva, L. Oliveira, and M. Pithon, ‘‘Automatic segmentating teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives,’’ Expert Syst. Appl., vol. 107, pp. 15–31, Oct. 2018.
[28] Y. Zhao et al., "Two-Stream Graph Convolutional Network for Intra-Oral Scanner Image Segmentation," in IEEE Transactions on Medical Imaging, vol. 41, no. 4, pp. 826-835, April 2022, doi: 10.1109/TMI.2021.3124217.
[29] https://www.researchgate.net/figure/Types-of-X-ray-images-a-Bitewing-X-ray-b-Periapical-X-ray-c-Panoramic-X-ray_fig1_330472299
[30] O. Oktay et al., “Attention u-net: Learning where to look for the pancreas,” ArXiv Prepr. ArXiv180403999, 2018.
[31] J. Schlemper et al., “Attention gated networks: Learning to leverage salient regions in medical images,” Med. Image Anal., vol. 53, pp. 197–207, 2019.
[32] A. Vaswani et al., “Attention is all you need,” 2017, pp. 5998–6008.
[33] Z. Zhang, H. Fu, H. Dai, J. Shen, Y. Pang, and L. Shao, “ET-Net: A Generic Edge-aTtention Guidance Network for Medical Image Segmentation,” 2019, pp. 442–450.
[34] Z. Si, D. Fu, and J. Li, “U-Net with Attention Mechanism for Retinal Vessel Segmentation,” in Image and Graphics, Cham, 2019, pp. 668–677.
[35] L. Mou et al., “CS-Net: Channel and Spatial Attention Network for Curvilinear Structure Segmentation,” in Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Cham, 2019, pp. 721–730.
[36] S. Lian, Z. Luo, Z. Zhong, X. Lin, S. Su, and S. Li, “Attention guided U-Net for accurate iris segmentation,” J. Vis. Commun. Image Represent., vol. 56, pp. 296–304, 2018.
[37] C. Wu, Y. Zou, and Z. Yang, “U-GAN: Generative Adversarial Networks with U-Net for Retinal Vessel Segmentation,” in 2019 14th International Conference on Computer Science & Education (ICCSE), Aug. 2019, pp. 642–646, doi: 10.1109/ICCSE.2019.8845397.
[38] N. Abraham and N. M. Khan, “A novel focal tversky loss function with improved attention u-net for lesion segmentation,” 2019, pp. 683–687.
[39] H. Zhang, H. Zhu, and X. Ling, “Polar coordinate sampling-based segmentation of overlapping cervical cells using attention U-Net and random walk,” Neurocomputing, vol. 383, pp. 212–223, 2020.
[40] Z. Fang, Y. Chen, D. Nie, W. Lin, and D. Shen, “RCA-U-Net: Residual Channel Attention U-Net for Fast Tissue Quantification in Magnetic Resonance Fingerprinting,” 2019, pp. 101–109.